



Deteksi Pornografi pada Karakter Animasi 2D dengan KNN (K-Nearest Neighbors) Menggunakan Fitur HSV

Pornography Detection of 2D Animated Characters with KNN (K-Nearest Neighbors) Using HSV Features

^{1)*} Candra Nur Mayasari, ²⁾ M. Arief Soeleman, ³⁾ Pujiono

^{1,2,3} Magister Teknik Informatika Universitas Dian Nuswantoro, Indonesia

Email: ^{1)} cnmayasari@gmail.com, ²⁾ arief22208@gmail.com, ³⁾ pujiono@dsn.dinus.ac.id

*Correspondence: ¹⁾ Candra Nur Mayasari

DOI:

10.36418/comserva.v2i08.462

Histori Artikel

Diajukan : 25-11-2022

Diterima : 05-12-2022

Diterbitkan : 10-12-2022

ABSTRAK

Perkembangan teknologi animasi sangat pesat, baik dalam bentuk 2D maupun 3D. Sebagian besar animator sering membuat karakter animasi perempuan untuk berbagai macam bidang seperti game, iklan atau anime lainnya. Seiring dengan perkembangan animasi juga terdapat dampak negatif dan positif, dimana dampak negatifnya adalah adanya simbol-simbol yang mengarah ke unsur pornografi. Kebanyakan animasi yang mengandung simbol negatif tersebut tersebar melalui internet, yang bisa diakses dengan mudah oleh siapa saja tidak pandang umur. Yang dapat berakibat kecanduan melihat pornografi sampai berperilaku negatif lainnya. Kecerdasan buatan yang saat berkembang pesat juga memungkinkan penelitian ini bertujuan untuk mendeteksi secara dini karakter animasi yang mengandung unsur pornografi. Dengan cara melakukan klasifikasi pada animasi 2D perempuan menggunakan metode algoritma KNN dengan ekstraksi fitur HSV. Fitur HSV cukup baik dalam mendeteksi warna-warna kompleks yang ada di karakter animasi 2D. Dimana dengan fitur HSV tersebut dapat membedakan warna kulit dengan warna lain yang ada pada animasi 2D. Dataset yang digunakan terdiri dari 3 kategori yaitu citra pornografi, citra semi pornografi, citra non pornografi. Dengan metode KNN dapat dilakukan klasifikasi antara gambar pornografi, semi pronografi, non pornografi. Dalam penelitian didapatkan hasil ekstraksi fitur HSV dengan metode klasifikasi KNN didapatkan nilai akurasi tertinggi adalah 63,16%.

Kata kunci: Animasi; Pornografi; Kecerdasan buatan; Ekstraksi Fitur; HSV; KNN

ABSTRACT

The development of animation technology is very rapid, both in 2D and 3D forms. Most animators often create female animated characters for various fields such as games, commercials or other anime. Along with the development of animation, there are also negative and positive impacts, where the negative impact is the presence of symbols that lead to pornography. Most of the animations containing negative symbols are spread through the internet, which can be accessed easily by anyone, regardless of age. Which can result in addiction to viewing pornography to other negative behaviors. Artificial intelligence, which is currently developing rapidly, also allows this research to aim at early detection of animated characters that contain pornographic elements. By classifying 2D female animation using the KNN algorithm method with HSV feature extraction. The HSV feature is quite good at detecting complex colors in 2D animated characters. Where the HSV feature can distinguish skin color from other colors in 2D animation. The dataset used consists of 3

categories, namely pornographic images, semi-pornographic images, and non-pornographic images. Using the KNN method, it is possible to classify pornographic, semi-pornographic and non-pornographic images. In this study, the results of HSV feature extraction using the KNN classification method obtained the highest accuracy value of 63.16%.

Keywords: Animation; Pornography; Artificial intelligence; Feature Extraction; HSV; KNN

PENDAHULUAN

Dengan berkembangnya teknologi, karakter animasi dibuat dengan berbagai macam jenis (Putra, 2019). Mulai dari karakter manusia hingga karakter hewan yang lucu dan menarik. Pada awalnya animasi dibuat untuk kalangan anak-anak, untuk memberi hiburan sekaligus pembelajaran yang disuguhkan dalam bentuk animasi yang menarik. Namun dengan seiring waktu karakter animasi mulai dibuat tidak hanya untuk anak-anak saja namun untuk semua kalangan umur. Di beberapa negara pembuatan karakter animasi sudah tidak hanya untuk film anak-anak saja, namun sudah untuk mempercantik website agar menarik, game baik game offline maupun game online (Mauludi, 2020). Seiring dengan perkembangan animasi juga terdapat dampak negatif dan positif, dimana dampak negatifnya adalah adanya simbol-simbol yang mengarah ke unsur pornografi (Ginting et al., 2021). Mulai dari penggunaan kostum yang kurang pantas, memperlihatkan bagian tubuh yang intim, memperlihatkan gambar yang vulgar dengan kostum yang mini. Kebanyakan animasi yang mengandung simbol negatif tersebut tersebar melalui internet, yang bisa diakses dengan mudah oleh siapa saja tidak pandang umur (Efendi, 2021). Yang dapat berakibat kecanduan melihat pornografi sampai berperilaku negatif lainnya. Oleh karena itu perlu adanya deteksi gambar animasi untuk mengendalikan gambar pornografi agar tidak dapat diakses dengan mudah terutama untuk anak-anak dibawah umur. Pengenalan kategori gambar pornografi, semi pornografi atau non pornografi perlu dilakukan untuk mempermudah dalam klasifikasi kategori gambar animasi tersebut (Hunaepi et al., 2019).

Pada penelitian tentang deteksi wajah dan pengenalan wajah pada kartun yang dilakukan oleh (Takayama et al., 2012) dengan menggunakan ekstraksi fitur yaitu warna kulit, garis tepi, kontur wajah dan kesimetrisan antara wajah dan rambut. Pada penelitian tersebut dilakukan segmentasi pada gambar karakter kartun dengan memisahkan warna kulit dengan bagian yang lain seperti rambut, mata dan baju. Hasil dari penelitian tersebut adalah Precision 0.476 recall 0.563 F-measure 0.516.

Penelitian tentang deteksi warna kulit pada karakter animasi dengan membandingkan 4 metode yang dilakukan Kazi Tanvir Ahmed Siddiqui dan Abu Wasif. Yang pertama dilakukan adalah dengan menggunakan metode deteksi warna kulit manusia dengan menggunakan algoritma RGB. Metode kedua adalah melakukan deteksi warna kulit pada karakter kartun dengan menggunakan hasil warna HSV. Dengan membandingkan 4 metode dari paper sebelumnya didapatkan hasil TP (True Positive) tertinggi yaitu 98.23% dan FP (False Positive) terendah yaitu 7.7 %. Warna kulit manusia memang mempunyai banyak warna dan terdapat bayangan yang terkadang dapat dideteksi sebagai bukan warna kulit sehingga deteksi warna kulit dengan menggunakan RGB. Sedangkan karakter animasi juga sangat beragam hampir sama dengan manusia, beragam dan terkadang antara karakter satu dengan karakter yang lain tidak yang sama warnanya. Pencampuran warna dalam pembuatan karakter animasi

memang sangat luar biasa banyak variannya sesuai dengan keinginan animator, sehingga cukup sulit dalam deteksi warna kulit pada karakter animasi.

Penelitian tentang deteksi dan klasifikasi berdasarkan warna kulit dengan menggunakan HSV dilakukan oleh (Areni et al., 2019). Penelitian ini untuk mengelompokkan warna kulit manusia, dimana dengan menggunakan sampel warna kulit orang Asia dan Eropa. Karena warna kulit orang Asia sangat beragam dan sesuai dengan ruang lingkup warna HSV. Dari hasil deteksi warna kulit tersebut dapat digunakan untuk melakukan filter gambar yang mengandung unsur pornografi dan non pornografi. Dan hasil nilai akurasi sebesar 90%. HSV mengandung unsur warna yang cukup kompleks dan sesuai dengan beragam warna kulit yang ada di Asia dan Eropa. Ada yang mempunyai warna terlalu terang maupun warna kulit yang terlalu gelap. Kemudian dapat membedakan warna kulit dengan rambut yang terkadang hampir mirip warnanya terutama pada orang Eropa yang mempunyai warna rambut kuning muda sampai dengan warna rambut yang gelap.

Penelitian tentang klasifikasi warna kulit dengan menggunakan rasio RGB dilakukan oleh (Yanuangga & Zaman, 2015). Pada penelitian ini dilakukan perbandingan beberapa metode untuk deteksi warna kulit. Setiap piksel yang menyusun warna kulit akan dihitung dan diukur untuk menentukan apakah warna kulit atau bukan. Hasil dari penelitian tersebut adalah TP (True Positive) 94.91 dan FP (False Positive) 33.07. Kelebihan menggunakan RGB adalah dapat mendeteksi warna-warna seperti warna bayangan pada kulit, warna kulit yang terlalu gelap atau warna kulit yang terlalu terang, kemudian komposisi warna yang menyusun dapat diambil sampelnya dan kemudian diuji.

Dengan melihat penelitian sebelumnya yang dapat digunakan sebagai state of art maka pada penelitian deteksi gambar porno karakter animasi 2D akan digunakan dengan menggunakan fitur warna HSV dan metode K-NN. Dari penelitian sebelumnya fitur warna HSV didapat nilai True Positive yang tinggi, serta variasi warna animasi yang banyak sehingga dipilihlah HSV sebagai fitur warna untuk deteksi pada penelitian kali ini. Sedangkan algoritma K-NN dipilih untuk memudahkan dalam melakukan klasifikasi sesuai dengan kategori pornografi, semi pornografi atau non pornografi (Febriyanti, 2020).

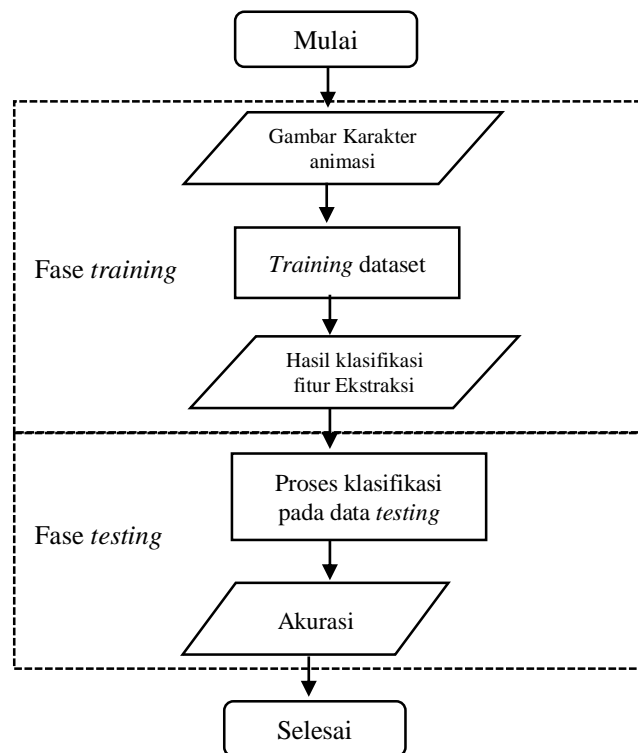
METODE

Penelitian dilakukan dengan pencarian data berupa gambar di internet dengan berbagai macam keyword yang berhubungan dengan karakter animasi. Kemudian gambar tersebut disimpan dengan folder masing-masing sesuai dengan jenis karakternya. Adapun gambar yang termasuk porno adalah gambar yang mempunyai banyak unsur warna kulit pada karakter animasi tersebut dan terdapat bagian tubuh vital yang terlihat.

Semua gambar dalam penelitian ini disimpan dalam format file .jpg dan berwarna cerah dengan piksel yang tidak terlalu kecil. Dataset yang akan digunakan sebanyak 300 gambar yang terdiri dari gambar karakter yang terdapat unsur porno sampai dengan gambar yang tidak ada unsur porno seperti gambar animasi hewan atau pemandangan. Gambar yang sudah didapat kemudian dilakukan pengecekan kembali dan dilakukan cropping serta resizing, namun tidak sampai gambar tersebut terlihat pecah dan tidak jelas. Proses tersebut untuk mempercepat komputasi pada pengolahan gambar karakter. Pada pemilahan gambar disebut dengan labelling. Dimana akan dibedakan gambar tersebut termasuk gambar porno atau bukan porno, agar nantinya memudahkan untuk dilakukan penelitian serta pemilahan gambar yang akan masuk ke dalam kategori data training dan gambar dalam kategori data testing.

Penelitian ini menggunakan dua tahapan yaitu tahap satu melakukan pengolahan dataset training dan tahap dua pengolahan dataset testing. Pada fase training dimulai dari gambar training dataset yang diperoleh dari gambar dataset asli kemudian dilakukan klasifikasi dengan menggunakan fitur ekstraksi. Dari ekstraksi tersebut didapatkan hasil ekstraksi sesuai dengan yang diharapkan. Selanjutnya masuk pada fase testing, data pada fase ini diperoleh dari hasil ekstraksi fase training. Hasil ekstraksi tersebut, kemudian dilakukan klasifikasi sesuai dengan metode yang sudah ditentukan. Hasilnya diperoleh nilai akurasi sesuai dengan metode algoritma yang telah dipilih. Berikut gambaran alur penelitian.

Pada proses fase training, dilakukan preprocessing pada gambar karakter animasi data training. Kemudian gambar tersebut dikonversi dari gambar RGB ke gambar HSV. Histogram diperoleh setelah proses tersebut selesai. Setelah didapat nilai warna dari proses tersebut, selanjutnya diklasifikasikan dengan metode K-NN, sehingga dihasilkan nilai parameter dari metode tersebut.



Gambar 1. Diagram Alur Penelitian

HASIL DAN PEMBAHASAN

Preprocessing Gambar

Preprocessing adalah tahapan awal dalam mengolah citra digital kemudian diidentifikasi ke tahap berikutnya (Zaenudin & Novamizanti, 2017). *Preprocessing* gambar karakter pada penelitian ini ada beberapa tahapan yaitu *resize* gambar, konversi citra ke *grayscale*, deteksi tepi, dilasi, erosi, dan kontur.

1. *Resize* Gambar

Resize gambar karakter dilakukan agar dalam pengolahan gambar karakter hasilnya tidak jauh berbeda *size* atau ukuran pada setiap *image*. Proses ini dilakukan agar ukuran gambar mempunyai *size*

yang sama atau seragam sehingga tidak berpengaruh pada hasil piksel setiap *image* yang digunakan dalam penelitian.



Gambar Asli



Gambar Hasil *Resize* (100,100)

Gambar 2. *Resize* Gambar Karakter

2. Konversi Citra ke *Grayscale*

Konversi citra ke *grayscale* adalah merubah warna citra asli menjadi citra keabuan atau *grayscale* (Saifullah et al., 2016). Proses ini digunakan untuk mendapatkan nilai intensitas warna antara 0 sampai dengan 255. Nilai 0 dalam *grayscale* menyatakan warna hitam dan nilai 255 dalam *grayscale* menyatakan warna putih.

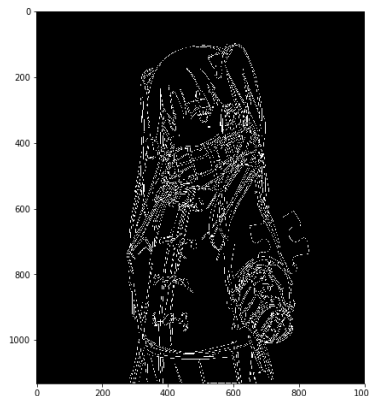
Dari proses konversi tersebut akan didapatkan nilai biner dari *image* yang akan diolah setiap pikselnya untuk penelitian. Nilai biner terdiri dari angka 0 (nol) dan angka 1 (satu). Nilai 0 (nol) merupakan warna hitam dan nilai 1 (satu) merupakan warna putih.



Gambar 3. Citra *Grayscale*

3. Deteksi Tepi

Proses ini dilakukan untuk memperoleh bagian tepi yang akan dilakukan seleksi dan mempermudah proses seleksi yang akan diambil untuk pemilihan bagian tertentu. Tujuannya untuk mengurangi distorsi pada *image*. Selanjutnya penghalusan pada bagian derau agar diperoleh bagian tepi yang kuat. Hasilnya, bagian tepi gambar diberi garis halus seperti garis putus-putus tipis untuk menandai bagian mana yang akan dipilih. Jenis metode deteksi tepi yang digunakan pada penelitian ini adalah *Canny Edge Detection*.



Gambar 4. Citra hasil *Canny Edge Detection*

4. Dilasi

Dilasi digunakan untuk mendapatkan efek suatu perbesaran dari dari segmen suatu obyek dimana obyek tersebut sudah berupa biner (Rani, 2020).

5. Erosi

Proses erosi adalah kebalikan dari proses dilasi. Erosi dilakukan untuk memperoleh nilai penyempitan dari suatu obyek biner (Sutariawan, 2018).

6. Kontur

Kontur adalah cara untuk menghasilkan tepi suatu obyek (Andarinny et al., 2017). Dengan menggunakan kontur akan didapatkan tepi suatu obyek yang digunakan dalam proses pengolahan citra. Kontur juga digunakan untuk memilih bagian obyek tertentu yang akan diolah dari obyek tersebut.

Ekstraksi Fitur

Data yang akan diolah awalnya berbentuk gambar dengan format jpg yang mengandung unsur warna RGB. Selanjutnya dari gambar tersebut akan dilakukan segmentasi citra dimana memisahkan gambar utama dengan gambar *background* yang tidak digunakan. Gambar karakter kemudian diolah dengan mengambil sampel warna kulit dari karakter animasi yang digunakan dalam data *training* dengan teknik warna HSV. Kemudian dilakukan ekstraksi warna pada gambar tersebut. Apabila pada gambar tersebut mengandung warna kulit, maka warna akan diubah menjadi warna putih. Sedangkan warna selain warna kulit akan diubah menjadi warna hitam. Dari hasil ekstraksi warna tersebut dapat dibedakan gambar yang mengandung warna kulit dan bukan warna kulit dengan diwakili dua (2) warna saja yaitu hitam dan putih. Kemudian jumlah piksel warna putih dan warna hitam dibandingkan. Apabila warna putih <20% maka gambar karakter animasi tersebut bukan termasuk pornografi. Dan apabila gambar mempunyai warna putih 20-50% artinya gambar tersebut merupakan semi pornografi, dan apabila >50% maka akan digolongkan menjadi gambar pornografi. Dengan langkah pengolahan data tersebut, akan dapat digunakan untuk mendeteksi gambar karakter animasi tersebut termasuk gambar pornografi atau gambar bukan pornografi.



Gambar 5. Citra Ekstraksi Fitur HSV

Untuk proses konversi RGB ke HSV dilakukan dengan mencari nilai S dan V menggunakan persamaan (1) sampai dengan persamaan (8).

Untuk mendapatkan nilai H, S, V berdasarkan R, G, B, terdapat beberapa cara. Cara yang paling sederhana adalah seperti berikut.

$$H = \tan\left(\frac{3(G-B)}{(R-G)+(R-B)}\right) \quad (1)$$

$$S = 1 - \frac{\min(R,G,B)}{V} \quad (2)$$

$$V = \frac{R+G+B}{3} \quad (3)$$

Namun cara ini membuat *hue* tidak terdefiniskan kalau S bernilai nol.

Cara kedua terdapat pada rumus yang digunakan seperti berikut.

$$r = \frac{R}{(R+G+B)}, g = \frac{G}{(R+G+B)}, b = \frac{B}{(R+G+B)} \quad (4)$$

$$V = \max(r, g, b) \quad (5)$$

$$S = \begin{cases} 0, & \text{jika } V = 0 \\ 1 - \frac{\min(r,g,b)}{V}, & V > 0 \end{cases} \quad (6)$$

$$H = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60*(g-b)}{S*V}, & \text{jika } V = r \\ 60 * \left[2 + \frac{b-r}{S*V}\right], & \text{jika } V = g \\ 60 * \left[4 + \frac{r-g}{S*V}\right], & \text{jika } V = b \end{cases} \quad (7)$$

$$H = H + 360 \text{ jika } H < 0 \quad (8)$$

Klasifikasi Data

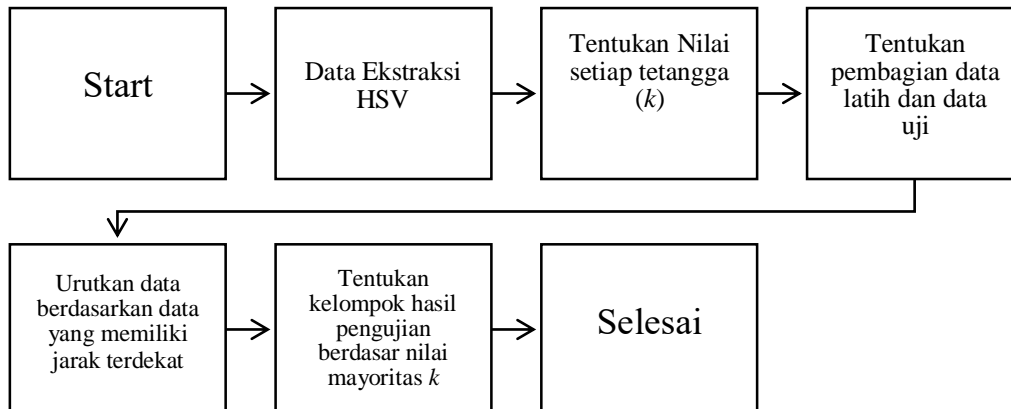
Data yang telah diolah dan sudah menemukan nilai dari HSV citra dari setiap data normal, kemudian dilakukan klasifikasi data sesuai labelnya. Dataset akan diolah dengan menggunakan *tools* Python. Sedangkan data *training* diolah dengan ekstraksi fitur warna HSV dan dikolaborasikan dengan algoritma K-NN.

Algoritma K-NN (*k Nearest Neighbor*) akan melakukan klasifikasi data citra uji ke dalam kelas yang mempunyai anggota banyak (Budianita et al., 2015). K-NN tersebut mempunyai prinsip kerja mencari jarak terdekat dari data yang akan diuji dengan k sebagai tetangga terdekat pada data latih. Data latih akan diproyeksikan ke ruang dimensi banyak, dan disetiap masing- masing dimensi akan merepresentasikan fitur datanya. Data latih yang sudah diklasifikasikan menjadi beberapa bagian sesuai dengan ruang dimensinya. Dimana setiap ruang terdapat titik yang akan dicari jarak terdekatnya dengan

k buah tetangganya. Jarak dekat dan jauh ketetanggaannya biasanya dihitung berdasarkan rumus *Euclidean* dengan rumus sebagai berikut:

$$dis(t_i, t_j) = \sqrt{\sum_{h=1}^k (t_{ih} - t_{jh})^2}$$

Sedangkan alur proses algoritma K-NN adalah sebagai berikut:



Gambar 6. Alur Proses K-NN

Pengujian dan Analisis

Berdasarkan penelitian sebelumnya, untuk mendeteksi warna kulit pada gambar karakter animasi masih menggunakan fitur ekstraksi warna RGB karena warna obyek yang akan diekstraksi masih mengandung unsur RGB. Sedangkan untuk penelitian kali ini akan menggunakan fitur ekstraksi warna HSV karena cukup rentan dalam pendeteksian warna untuk dilakukan seleksi pada citra.

Tahap selanjutnya adalah melakukan konversi warna dari RGB ke HSV. Hasil konversi dari RGB ke HSV akan dibuat dalam grafik histogram kemudian hasilnya akan dilakukan perbandingan dengan *database* yang ada. Pada tahap ini untuk mengetahui nilai dari konversi RGB ke HSV maka perlu dilakukan perhitungan konversi RGB ke HSV. Dimana kita mencari nilai dari *Hue*, *Saturation* dan *Value*. *Hue* akan mempresentasikan warna asli dari citra tersebut, *Saturation* akan memperlihatkan efek warna putih yang akan menunjukkan tingkat warna yang mendominasi, *Value* menunjukkan perbedaan dari warna aslinya dimana nanti warna putih akan mewakili warna kulit dan warna hitam adalah obyek selain warna kulit. Untuk obyek yang mempunyai banyak warna maka perlu dihitung untuk nilai H dan nilai S.

Pada ekstraksi fitur HSV terlebih dahulu akan dicari nilai *Hue* dengan ditentukan nilai maksimal dari R, G, B. Jika dilihat warna dominan ke merah, maka rumusnya:

$$\text{Hue} = \frac{G-B}{\text{RGBmax}-\text{RGB min}} \times 60$$

Bila dominan hijau, maka rumusnya:

$$\text{Hue} = 120 + \frac{B-R}{\text{RGBmax}-\text{RGB min}} \times 60$$

Bila dominan biru, maka rumusnya:

$$\text{Hue} = 240 + \frac{R-G}{\text{RGBmax}-\text{RGB min}} \times 60$$

Jika dijabarkan adalah sebagai berikut:

$$r = \frac{R}{((1+x)^n = 1 + \frac{nx}{1} + \frac{n(n-1)x^2}{2!} + \dots + R+G+B)}, g = \frac{G}{(R+G+B)}, b = \frac{B}{(R+G+B)}$$

$$V = \max(r, g, b)$$

$$S = \begin{cases} 0, & \text{jika } V = 0 \\ 1 - \frac{\min(r,g,b)}{V}, & V > 0 \end{cases}$$

$$H = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60 * (g-b)}{S * V}, & \text{jika } V = r \\ 60 * \left[2 + \frac{b-r}{S * V} \right], & \text{jika } V = g \\ 60 * \left[4 + \frac{r-g}{S * V} \right], & \text{jika } V = b \end{cases}$$

$$H = H + 360 \text{ jika } H < 0$$

Pola dari penyebaran warna kulit pada tiap citra adalah berbeda-beda saat dilakukan analisis *Hue, Saturation, Value*. Agar deteksi kulit dapat terfasilitasi maka ditentukan nilai tertentu untuk warna kulit. Dengan memberikan ketentuan nilai bila warna kulit dibawah 20% maka tergolong non porno, apabila antara 20% - 50% maka tergolong semi porno, dan jika diatas 50% maka tergolong porno.

Dalam penelitian kali ini citra akan dibuat matriks dengan ordo 4 x 4 untuk semua citra dataset setelah dilakukan *resize*. Selanjutnya dihitung konversi dari RGB ke HSV dengan rumus yang telah ada. Proses berikutnya adalah merubah citra yang tadinya *grayscale* menjadi citra biner. Kemudian setelah didapatkan citra biner maka langkah selanjutnya dilakukan proses morfologi agar mendapatkan citra yang lebih baik dan lebih signifikan pada bagian warna yang akan diseleksi nantinya.

HSV adalah salah satu segmentasi citra dengan menggunakan sebuah algoritma tertentu. Kemudian tiap *pixel* yang dihasilkan akan membentuk matriks yang dapat digunakan untuk mencari probabilitas suatu kedekatan dari dua citra yaitu citra pembandingan dan *sample* citranya. Perhitungan dari probabilitasnya dengan rumus:

$$P(y) = \frac{1}{\sqrt{(2\pi)^3 |C|}} \exp \left(-\frac{1}{2} (x - \bar{x})^T C^{-1} (x - \bar{x}) \right)$$

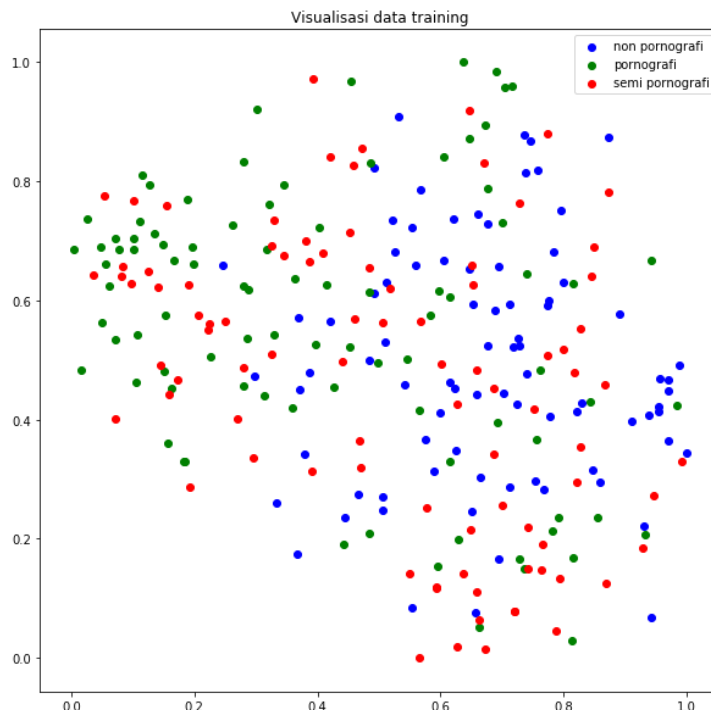
Penggunaan rumus diatas untuk menghitung kemungkinan nilai dengan range 0-1, bila nilai mendekati 1 maka akan diberi warna putih, bila nilai mendekati 0 maka akan diberi warna hitam. Dari penelitian yang telah dilakukan pada semua citra diperoleh data sampel setelah dilakukan ekstraksi fitur H, S, V pada citra seluruh dataset.

Tabel 1. Tabel Sampel Ekstraksi Fitur HSV

<pre> [[0.4832, 0., 0.0048, ..., 0.0128, 0.0368, 0.8944], [0.8576, 0., 0., ..., 0.0032, 0.0352, 0.3264], [0.7632, 0., 0., ..., 0.0176, 0.08, 0.5472], ..., [1., 0., 0., ..., 0.024, 0.0832, 0.1568], [0.872, 0.0048, 0.0064, ..., 0.0576, 0.1632, 0.2144], [0.5296, 0.0032, 0., ..., 0.008, 0.0256, 0.0112]] </pre>

Pada data sampel di atas menunjukkan hasil konversi dari RGB ke HSV setelah dilakukan perhitungan. Dari hasil tersebut akan diperoleh hasil seleksi warna berdasarkan warna kulit dengan fitur H, S, V. Dan didapatkan hasil data untuk data *training* dan data *testing* yang jumlahnya sama dengan jumlah data aslinya dimana 100 gambar non pornografi, 100 semi pornografi, 100 pornografi, dimana semuanya sudah diberi label. Untuk data *testing* kita masukkan 16 gambar non pornografi, 11 gambar semi pornografi, 11 gambar pornografi dengan total data *testing* 38 citra. Sedangkan data *training* yang digunakan adalah 84 non pornografi, 89 pornografi dan 89 semi pornografi dengan total data *training* 262 citra.

Dari hasil ekstraksi fitur yang telah dilakukan yaitu ekstraksi H, S, V pada 262 citra *training* maka dihasilkan 12287 fitur per citra. Dimana dalam melakukan ekstraksi HSV ini juga dengan menggunakan tekstur haralick untuk mengetahui tingkat pornografinya. Sehingga didapatkan visualisasi penyebaran dari data *training* seperti terlihat dalam gambar dibawah ini. Sehingga cukup banyak kolom yang terbentuk untuk setiap gambarnya dalam menentukan label datanya.



Gambar 7. Visualisasi Seluruh Data *Training* HSV

Pengujian Algoritma k-Nearest Neighbor (K-NN)

Setelah ekstraksi fitur ditentukan, berikutnya dilakukan pemrosesan algoritma K-NN. Hasil pengujian akan sangat variatif yang akan dituangkan dalam tabel dari hasil perhitungan algoritma K-NN. Dimana metode K-NN adalah mencari sejumlah k pola dari data *training* dengan data *input*, kemudian menentukan kelas keputusan berdasarkan tetangga terdekatnya. Maka akan diperoleh nilai k yang terbaik yang mana untuk mendapatkan nilai akurasi yang terbaik dari nilai k yang telah dilakukan pengujian. Yang pertama dilakukan melakukan pengujian pada data *training* untuk mendapatkan pola latih kemudian disimpan. Setelah proses pengujian pada data *training* dilakukan tahap selanjutnya adalah proses pengujian pada data *testing* dengan memasukkan sejumlah nilai k untuk memperoleh pola terdekat.

Tabel 2. Hasil Klasifikasi K-NN Terhadap Citra *Training*

	Data	Distance	Ranking	Label
94	d95, dx	2.4867	1	Non pornografi
116	d117, dx	2.5200	2	Semi pornografi
246	d247, dx	2.5329	3	Non pornografi
105	d106, dx	2.6493	4	Non pornografi
211	d212, dx	2.6671	5	Non pornografi
158	d159, dx	2.9299	27	Pornografi
...
43	d44, dx	4.0142	233	Semi pornografi
30	d31, dx	4.0228	234	Pornografi

Tabel 2. di atas merupakan perhitungan dari *Euclidean Distance*, dimana nilai k yang diambil adalah dari k = 1 dan seterusnya. Maka dari hasil perhitungan akan didapatkan nilai jarak yang terbaik dan terdekat, sehingga akan dapat diperoleh nilai akurasi yang terbaik. Data tersebut merupakan termasuk kelas mayoritas yang terbentuk sesuai label yang telah ditentukan fiturnya.

Prediksi →

Aktual ↓	porno	69	31	5
	semi	0	63	26
	non	4	36	28
		porno	semi	non

Gambar 8. *Confusion Matrix* Metode K-NN Data *Training*

Hasil *confusion matrix* untuk data *training* maka diperoleh untuk data citra porno 69 gambar, semi porno 63 gambar dan non porno sebanyak 28 gambar. Sehingga dapat diperoleh nilai akurasinya

dengan menghitung nilai TP (*True False*) dibagi dengan jumlah semua datasetnya seperti pada rumus berikut:

$$\text{Akurasi} = \frac{TP}{\text{Dataset}} \times 100\%$$

$$\text{Akurasi} = \frac{69+63+28}{262} \times 100\%$$

$$\text{Akurasi} = 61\%$$

Jadi diperoleh nilai akurasi untuk data *training* dari *confusion matrix* tersebut adalah sebesar 61%.

Akurasi Pengujian

Untuk hasil ekstraksi fitur data pada data *testing* dengan jumlah 16 gambar non pornografi, 11 gambar semi pornografi dan 11 gambar pornografi dengan jumlah 38 citra data *testing*. Tabel berikut merupakan salah satu contoh hasil ekstraksi fitur data *testing* yang akan dicari kelasnya. Dimana label hasil ekstraksi menunjukkan citra tersebut termasuk dalam non pornografi.

Tabel 3. Hasil Salah Satu Fitur Dari Data Testing

Label	F0	F1	F2	F3	...	F12283	F12284	F12285	F12286	F12287
1 Non pornografi	1.0	0.0	0.0	0.0	...	0.0	0.0032	0.0016	0.0336	0.6704

Setelah dicari nilai jarak *Euclidean Distance* sesuai perhitungan dan dilakukan perankingan berdasarkan jarak terdekat dan didapatkan label yang sesuai, maka hasil klasifikasi akan didapatkan nilai akurasi yang terbaik. Dimana nilai k yang dipakai adalah k = 1 sampai k = 10.

Hasil ekstraksi fitur pada data testing dengan jumlah 38 gambar dapat dilihat pada *confusion matrix* pada gambar berikut.

		Prediksi →		
Aktual ↓	porno	7	2	5
	semi	3	10	0
	non	2	2	7
		porno	semi	non

Gambar 9. *Confusion Matrix* Metode K-NN Data Testing

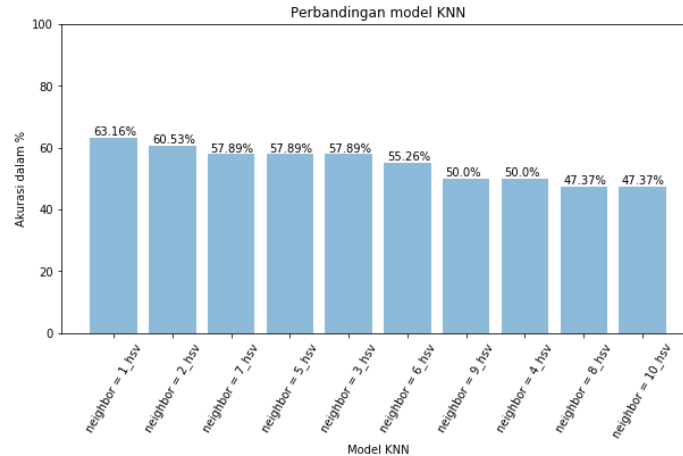
Dari data *confusion matrix* tersebut diketahui bahwa data citra porno yang diprediksi porno dan benar citra porno sebanyak 7 gambar, data yang diprediksi semi porno dan benar citra semi porno sebanyak 10 citra, sedangkan citra yang diprediksi non pornografi dan benar citra non pornografi sebanyak 7 citra. Total data *testing* yang digunakan sebanyak 38 gambar. Perhitungan nilai akurasinya seperti rumus berikut:

$$\text{Akurasi} = \frac{TP}{\text{Dataset}} \times 100\%$$

$$\text{Akurasi} = \frac{7+10+7}{38} \times 100\%$$

$$\text{Akurasi} = 63,16\%$$

Perbandingan hasil akurasi dari model K-NN dapat dilihat pada gambar berikut.



Gambar 10. Perbandingan Nilai Akurasi Model K-NN

Perbandingan nilai akurasi seperti pada gambar tersebut didapatkan dari hasil pemrosesan data *training* dan data *testing* dengan menggunakan Python. Hasil tersebut didapat dari ekstraksi fitur HSV dengan menggunakan algoritma K-NN yang diterapkan pada gambar animasi 2D. Dari pengujian algoritma K-NN didapatkan nilai akurasi dengan jarak terdekat yaitu dengan $k = 1$ dengan hasil akurasinya 63.16% dari total data *testing* 38 gambar. Nilai k yang digunakan pada hasil akurasi tersebut merupakan nilai akurasi terbaik dari pola k yang lain. Data tersebut juga bisa dilihat pada tabel berikut.

Tabel 4. Perbandingan Nilai Akurasi Fitur HSV Metode K-NN

No	Nama Metode	Ekstraksi Fitur	Pembeda	Hasil Akurasi
1	K-NN	HSV	k=1	63,16 %
2	K-NN	HSV	k=2	60,53 %
3	K-NN	HSV	k=3	57,89 %
4	K-NN	HSV	k=4	50 %
5	K-NN	HSV	k=5	57,89 %
6	K-NN	HSV	k=6	55,26 %
7	K-NN	HSV	k=7	57,89 %
8	K-NN	HSV	k=8	47,37 %
9	K-NN	HSV	k=9	50 %
10	K-NN	HSV	k=10	47,37 %

Nilai akurasi yang didapatkan dari pengujian dengan menggunakan ekstraksi fitur HSV dan metode algoritma K-NN seperti pada tabel 4. tersebut menampilkan beberapa perbandingan nilai akurasi. Pada pengujian didapatkan 10 data dimulai dari nilai $k=1$ sampai dengan $k=10$ yang semuanya

menggunakan metode K-NN dengan kombinasi ekstraksi fitur HSV pada citra karakter animasi 2D. Urutan dari tertinggi dari nilai akurasi terdapat pada k=1 yaitu sebesar 63,16 %. Sedangkan nilai akurasi tertinggi kedua terdapat pada k=2 dengan nilai 60,53%. Nilai akurasi pada k=3 sama dengan nilai k=5 dan nilai akurasi pada k=7 yaitu sebesar 57,89 %. Urutan selanjutnya terdapat pada k=6 sebesar 55,26 %. Sedangkan nilai k=4 sama dengan k=9 berada diurutan berikutnya dengan nilai sama sebesar 50 %. Dan urutan terakhir dari 10 data tersebut sebesar 47,37 % terdapat pada nilai k=8 dan nilai k=10.

SIMPULAN

Dari penelitian yang telah dilakukan untuk dapat menentukan gambar animasi 2D apakah tergolong porno, semi porno atau non porno dapat dilakukan dengan salah satu cara mencari warna kulit dengan melakukan ekstraksi fitur HSV. Dimana dari citra yang semua hasil download harus melalui proses preprocessing terlebih dahulu. Mulai dari proses segmentasi, proses cropping dan resize pada semua citra. Hal ini untuk memudahkan proses dalam ekstraksi fitur pada citra.

Ekstraksi fitur dimulai dari mengubah citra RGB menjadi HSV dengan mencari nilai H (Hue), S (Saturation), V (Value). Kemudian menentukan warna kulit dari citra animasi agar dapat ditentukan mana yang termasuk porno, semi porno dan non porno. Apabila warna kulit < 20% maka masuk label non porno, apabila warna kulit antara 20% – 50% maka masuk label semi porno dan apabila diatas 50% maka masuk label porno. Untuk metode klasifikasi peneliti menggunakan metode algoritma K-NN. Dari hasil perhitungan didapatkan nilai k = 1 mempunyai akurasi sebesar 63,16% dengan cara mencari jarak terdekat atau Euclidean Distance dengan memasukkan nilai k = 1-10.

Dari nilai akurasi yang diperoleh masih belum menunjukkan hasil yang sangat besar tetapi setidaknya klasifikasi dengan ekstraksi fitur HSV dan metode algoritma K-NN ini dapat melakukan deteksi pornografi pada gambar animasi 2D terutama untuk gambar animasi perempuan.

DAFTAR PUSTAKA

- Andarinny, A. A., Widodo, C. E., & Adi, K. (2017). Perancangan sistem identifikasi biometrik jari tangan menggunakan Laplacian of Gaussian dan ekstraksi kontur. *Youngster Physics Journal*, 6(4), 304–314.
- Areni, I. S., Amirullah, I., & Arifin, N. (2019). Klasifikasi Kematangan Stroberi Berbasis Segmentasi Warna dengan Metode HSV. *Jurnal Penelitian Enjiniring*, 23(2), 113–116.
- Budianita, E., Jasril, J., & Handayani, L. (2015). Implementasi Pengolahan Citra dan Klasifikasi K-Nearest Neighbour Untuk Membangun Aplikasi Pembeda Daging Sapi dan Babi Berbasis Web. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, 12(2), 242–247.
- Efendi, Z. (2021). *Budaya Pop Dan Persaingan Identitas (Studi Deskriptif Pada Komunitas Anime Attack On Titan)*. UMSU.
- Febriyanti, A. (2020). *Analisis Sentimen Persepsi Pengguna JNE Menggunakan Algoritma Naïve Bayes Classifier*.
- Ginting, R., Yulistiyono, A., Rauf, A., Manullang, S. O., Siahaan, A. L. S., Kussanti, D. P., PS, T. E. A., Djaya, T. R., Ayu, A. S., & Effendy, F. (2021). *Etika Komunikasi Dalam Media Sosial: Saring Sebelum Sharing* (Vol. 1). Penerbit Insania.
- Hunaepi, A., Makhsun, M., & Sarwani, S. (2019). Deteksi Situs Pornografi Berdasarkan Gambar Menggunakan Metode Convolutional Neural Network. *Jurnal Teknik Informatika Vol*, 12(2), 137.
- Mauludi, S. (2020). *Socrates Cafe-Bijak, Kritis & Inspiratif Seputar Dunia & Masyarakat Digital*. Elex Media Komputindo.
- Putra, G. L. A. K. (2019). Pemanfaatan Animasi Promosi Dalam Media Youtube. *SENADA (Seminar Nasional Manajemen, Desain Dan Aplikasi Bisnis Teknologi)*, 2, 259–265.
- Rani, S. (2020). *Sistem Identifikasi Sifat Fenotip Berdasarkan Citra Digital pada Buah Mentimun*.
- Saifullah, S., Sunardi, S., & Yudhana, A. (2016). Perbandingan segmentasi pada citra asli dan citra kompresi wavelet untuk identifikasi telur. *ILKOM Jurnal Ilmiah*, 8(3), 190–196.
- Sutariawan, I. P. E. (2018). Segmentasi Mata Katarak pada Citra Medis Menggunakan Metode Operasi Morfologi. *Jurnal Ilmu Komputer Indonesia*, 3(1), 23–31.
- Takayama, K., Johan, H., & Nishita, T. (2012). Face detection and face recognition of cartoon characters using feature extraction. *Image, Electronics and Visual Computing Workshop*, 48.
- Yanuangga, G. H. L., & Zaman, L. (2015). Deteksi Jerawat Otomatis Pada Citra Wajah Studi Kasus pada Kulit Penduduk Jawa. *Seminar Nasional Inovasi Dalam Desain Dan Teknologi-Ideatech 2015*.
- Zaenudin, C. A., & Novamizanti, L. (2017). Perancangan Alat Identifikasi Keaslian Dan Nominal Mata Uang Kertas Real Time Untuk Penyandang Tunanetra Berbasis Template Matching Dengan Raspberry Pi. *ReTII*.



© 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).